

12.2 Frequency and Histograms

Frequency: the number of data values in the interval

Frequency table: groups a set of data values into intervals and shows the frequency for each interval; intervals do not overlap, do not have any gaps, and are usually of equal size

Histogram: a graph that can display data from a frequency table; has one bar for each interval; the height of the bar shows the frequency of the data in the interval it represents; no gaps between bars; bars are usually of equal width

Uniform: bars are roughly the same height

Symmetric: a vertical line can divide the histogram into two parts that are close to mirror images

Skewed: has one peak that is NOT in the center

Cumulative frequency table: shows the number of data values that lie in or below a given interval

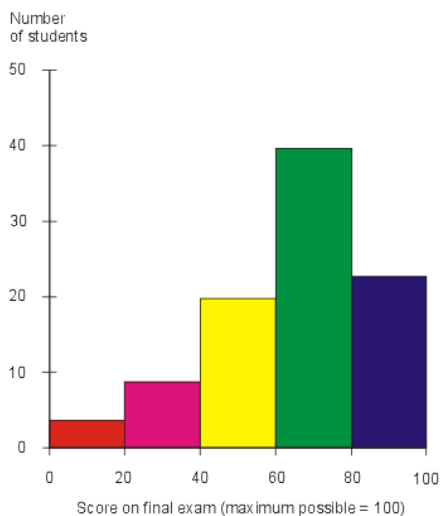
Examples

- The marks awarded for an assignment set for a Year 8 class of 20 students were as follows:

6 7 5 7 7 8 7 6 9 7
4 10 6 8 8 9 5 6 4 8

Mark	Tally	Frequency
4		2
5		2
6		4
7		5
8		4
9		2
10		1

- Histogram (skewed)



12.3 Measures of Central Tendency and Dispersion

Measure of Central Tendency: a way to describe the center of a data set

Mean: the average; $\frac{\text{sum of the data values}}{\text{number of data values}}$

Median: the middle value in a data set when the **values are arranged in ORDER**; when there are two middle values, the median is the average of these two middle data values

Mode: most occurring; the data item that occurs the **MOST** times; a data set can have one mode, no mode, or more than one mode

Outlier: a data value that is much greater or less than the other values in the set

Measure of dispersion: describes how dispersed, or spread out, the values in a data set are

Range of a set of data: a measure of dispersion; the difference between the greatest and the least data values

Example

Data: 2, 7, 12, 5, 4, 6, 10, 5, 9, 2, 7, 3

Mean: $\frac{2+7+12+5+4+6+10+5+9+2+7+3}{12} = \frac{72}{12} = 6$

Median: 2, 2, 3, 4, 5, 5, 6, 7, 7, 9, 10, 12 $\rightarrow \frac{5+6}{2} = 5.5$

Mode: 2, 5, 7

Range: $12 - 2 = 10$

12.4 Box-and-Whisker Plots

Quartile: value that divides a set into four equal parts

Second quartile, Q_2 : the median, which separates the data into upper and lower halves

First quartile, Q_1 : the median of the lower half of the data

Third quartile, Q_3 : the median of the upper half of the data

Interquartile range: the difference between the third and first quartiles

Box-and-whisker plot: a graph that summarizes a set of data by displaying it along a number line; consists of three parts: a box and two whiskers

Percentiles: separate data sets into 100 equal parts

Percentile rank: the percentage of data values that are less than or equal to that value

Examples:

Data: 2, 7, 12, 5, 4, 6, 10, 5, 9, 2, 7, 3

2, 2, 3, 4, 5, 5, 6, 7, 7, 9, 10, 12
 ↓ ↓ ↓
 Q_1 Q_2 Q_3

Minimum: 2

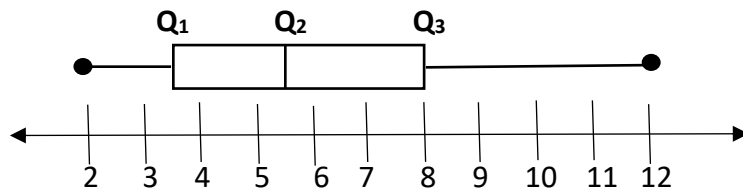
First Quartile: 3.5

Second Quartile: 5.5

Third Quartile: 8

Maximum: 12

Interquartile range: $8 - 3.5 = 4.5$



Percentile rank: 7 has a percentile rank of 75^{th} $\rightarrow \frac{9 \text{ numbers are equal to or less than 7 in the data}}{12 \text{ numbers are in the data}} = \frac{9}{12} = 0.75$